



THE UNIVERSITY of EDINBURGH

## Edinburgh Research Explorer

# Novel Associations of Multiple Genetic Loci With Plasma Levels of Factor VII, Factor VIII, and von Willebrand Factor The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium

### Citation for published version:

Wellcome Trust Case Control Consor, Smith, NL, Chen, M-H, Dehghan, A, Strachan, DP, Basu, S, Soranzo, N, Hayward, C, Rudan, I, Sabater-Lleal, M, Bis, JC, de Maat, MPM, Rumley, A, Kong, X, Yang, Q, Williams, FMK, Vitart, V, Campbell, H, Maelarstig, A, Wiggins, KL, Van Duijn, CM, McArdle, WL, Pankow, JS, Johnson, AD, Silveira, A, McKnight, B, Uitterlinden, AG, Aleksic, N, Meigs, JB, Peters, A, Koenig, W, Cushman, M, Kathiresan, S, Rotter, JI, Bovill, EG, Hofman, A, Boerwinkle, E, Tofler, GH, Peden, JF, Psaty, BM, Leebek, F, Folsom, AR, Larson, MG, Spector, TD, Wright, A, Wilson, JF, Hamsten, A, Lumley, T, Witteman, JCM, Tang, W & O'Donnell, CJ 2010, 'Novel Associations of Multiple Genetic Loci With Plasma Levels of Factor VII, Factor VIII, and von Willebrand Factor The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium', *Circulation*, vol. 121, no. 12, pp. 1382-U45. <https://doi.org/10.1161/CIRCULATIONAHA.109.869156>

### Digital Object Identifier (DOI):

[10.1161/CIRCULATIONAHA.109.869156](https://doi.org/10.1161/CIRCULATIONAHA.109.869156)

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Peer reviewed version

### Published In:

Circulation

### Publisher Rights Statement:

Published in final edited form as:  
Circulation. 2010 March 30; 121(12): 1382–1392. doi:10.1161/CIRCULATIONAHA.109.869156.

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



Published in final edited form as:

*Circulation*. 2010 March 30; 121(12): 1382–1392. doi:10.1161/CIRCULATIONAHA.109.869156.

## Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE Consortium

Nicholas L. Smith, PhD\*, Ming-Huei Chen, PhD\*, Abbas Dehghan, MD, DSc\*, David P. Strachan, MD\*, Saonli Basu, PhD\*, Nicole Soranzo, MSc, PhD, Caroline Hayward, PhD, Igor Rudan, PhD, Maria Sabater-Lleal, PhD, Joshua C. Bis, PhD, Moniek de Maat, PhD, Ann Rumley, PhD, Xiaoxiao Kong, MS, ME, Qiong Yang, PhD, Frances M. K. Williams, PhD, Veronique Vitart, PhD, Harry Campbell, MD, PhD, Anders Mälarstig, PhD, Kerri L. Wiggins, MS, RD, Cornelia Van Duijn, PhD, Wendy L. McArdle, PhD, James S. Pankow, PhD, MPH, Andrew D. Johnson, PhD, Angela Silveira, PhD, Barbara McKnight, PhD, Andre Uitterlinden, PhD, Wellcome Trust Case Control Consortium, Nena Aleksic, PhD, James B. Meigs, MD, MPH, Annette Peters, PhD, Wolfgang Koenig, MD, Mary Cushman, MD, MSc, Sekar Kathiresan, MD, Jerome I. Rotter, MD, Edwin G. Bovill, MD, Albert Hofman, MD, PhD, Eric Boerwinkle, PhD, Geoffrey H. Tofler, MD, John F. Peden, PhD, Bruce M. Psaty, MD, PhD, Frank Leebek, MD, PhD, Aaron R. Folsom, MD, MPH, Martin G. Larson, SD, Timothy D. Spector, MD, Alan F. Wright, MD, PhD, James F. Wilson, DPhil, Anders Hamsten, MD, PhD, Thomas Lumley, PhD\*\*, Jacqueline C. Witteman, PhD\*\*, Weihong Tang, MD, PhD\*\*, and Christopher J. O'Donnell, MD, MPH\*

ARIC: Divisions of Biostatistics (SB, XK), Epidemiology and Community Health (JSP, ARF, WT), University of Minnesota, Minneapolis, Minnesota, USA; DNA Laboratory (NA), Human Genetics Center and Institute of Molecular Medicine (EB), University of Texas Health Science Center at Houston, Houston, Texas, USA; B58C: Division of Community Health Sciences, St George's, University of London, UK (DPS); Division of Cardiovascular and Medical Sciences, University of Glasgow, Royal Infirmary, Glasgow, UK (AR); ALSPAC Laboratory, University of Bristol, UK (WLM); CHS: Departments of Epidemiology (NLS, BMP), Medicine (JCB, K LW, BMP), Biostatistics (TL, BM), Health Services (BMP), University of Washington; and Group Health Research Institute (BMP), Group Health Cooperative, Seattle, Washington, USA; Departments of Pathology (MC, EGB), Medicine (MC), University of Vermont, Burlington, Vermont USA; Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA (JIR); FHS: National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, MA, USA (MHC, QY, ADJ, SK, MGL, CJO); Departments of Biostatistics (QY), Neurology (MHC), and Mathematics (MGL), Boston University, Boston, MA, USA; Division of Intramural Research, National Heart, Lung and Blood Institute, Bethesda MD, USA (ADJ, CJO); General Medicine Division, Massachusetts General Hospital, Harvard Medical School, Boston MA, USA (JBM);

Corresponding authors: Nicholas L. Smith, PhD, Cardiovascular Health Research Unit, University of Washington, 1730 Minor Ave, Suite 1360, Seattle WA 98101, Phone: 206-287-2777; nlsmith@u.washington.edu; Christopher J. O'Donnell, MD, MPH, NHLBI's Framingham Heart Study, 73 Mount Wayte Avenue, Suite #2, Framingham, MA 01702, Phone: 508-935-3435; qdonnellc@nhlbi.nih.gov.

\*These authors contributed equally as first authors.

\*\*These authors contributed equally as last authors.

### Disclosures

No author has a real or perceived conflict of interest to report for this manuscript except James B. Meigs who reports receiving grant support from Sanofi-Aventis and GlaxoSmithKline, and serving as a consultant/advisor to Interleukin Genetics and Eli Lilly.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Cardiology Division, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston MA, USA (SK, CJO); Royal North Shore Hospital, Sydney, Australia (GHT); RS: Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands (AD, CVD, AH, AU, JCM), Department of Hematology, Erasmus Medical Center, Rotterdam, The Netherlands (MdM, FL), Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands (AU), Member of the Netherlands Consortium on Healthy Aging (AD, JCM); Twins UK: Wellcome Trust Sanger Institute, Hinxton, (NS), Department of Twin Research & Genetic Epidemiology, King's College London, London UK (NS, FMKW, TDS); VIS: MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Edinburgh, Scotland, UK (CH, VV, AFW), Croatian Centre for Global Health, University of Split Medical School, Soltanska, Croatia (IR), Centre for Population Health Sciences, University of Edinburgh Medical School, Teviot Place, Edinburgh, Scotland, UK (IR, HC, JFW); ORCADES: Centre for Population Health Sciences, University of Edinburgh Medical School, Teviot Place, Edinburgh, Scotland, UK (IR, HC, JFW).MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Edinburgh, Scotland, UK (CH, VV, AFW), Croatian Centre for Global Health, University of Split Medical School, Soltanska, Croatia (IR); PROCARDIS: Atherosclerosis Research Unit, Department of Medicine, Karolinska Institutet, Stockholm, Sweden (MSL, AM, AS, AH), Department of Cardiovascular Medicine, University of Oxford, Oxford, UK (JFP); KORA: Institute of Epidemiology, Helmholtz Zentrum München – German Research Center for Environment and Health, Neuherberg (AP); Internal Medicine II – Cardiology, University of Ulm, Ulm, Germany (WK)

## Abstract

**Background**—Plasma levels of coagulation factors VII (FVII), VIII (FVIII), and von Willebrand factor (vWF) influence risk of hemorrhage and thrombosis. We conducted genome-wide association studies to identify new loci associated with plasma levels.

**Methods and Results**—Setting includes 5 community-based studies for discovery comprising 23,608 European-ancestry participants: ARIC, CHS, B58C, FHS, and RS. All had genome-wide single nucleotide polymorphism (SNP) scans and at least 1 phenotype measured: FVII activity/antigen, FVIII activity, and vWF antigen. Each study used its genotype data to impute to HapMap SNPs and independently conducted association analyses of hemostasis measures using an additive genetic model. Study findings were combined by meta-analysis. Replication was conducted in 7,604 participants not in the discovery cohort. For FVII, 305 SNPs exceeded the genome-wide significance threshold of  $5.0 \times 10^{-8}$  and comprised 5 loci on 5 chromosomes: 2p23 (smallest p-value  $6.2 \times 10^{-24}$ ), 4q25 ( $3.6 \times 10^{-12}$ ), 11q12 ( $2.0 \times 10^{-10}$ ), 13q34 ( $9.0 \times 10^{-259}$ ), and 20q11.2 ( $5.7 \times 10^{-37}$ ). Loci were within or near genes, including 4 new candidate genes and *F7* (13q34). For vWF, 400 SNPs exceeded the threshold and marked 8 loci on 6 chromosomes: 6q24 ( $1.2 \times 10^{-22}$ ), 8p21 ( $1.3 \times 10^{-16}$ ), 9q34 ( $<5.0 \times 10^{-324}$ ), 12p13 ( $1.7 \times 10^{-32}$ ), 12q23 ( $7.3 \times 10^{-10}$ ), 12q24.3 ( $3.8 \times 10^{-11}$ ), 14q32 ( $2.3 \times 10^{-10}$ ) and 19p13.2 ( $1.3 \times 10^{-9}$ ). All loci were within genes, including 6 new candidate genes, as well as *ABO* (9q34) and *VWF* (12p13). For FVIII, 5 loci were identified and overlapped vWF findings. Nine of the 10 new findings replicated.

**Conclusions**—New genetic associations were discovered outside previously known biologic pathways and may point to novel prevention and treatment targets of hemostasis disorders.

## Keywords

genome-wide variation; factor VII; factor VIII; von Willebrand factor; epidemiology; meta-analysis; thrombosis; hemostasis

A complex cascade of coagulation factors underlies hemostasis and prevents life-threatening blood loss from damaged blood vessels. The hemostatic factors VII and VIII, both produced

in the liver, play central roles in the initiation and propagation, respectively, of fibrin formation. In the tissue-factor pathway, blood coagulation factor VII (FVII), once activated, serves as a catalyst for factor X (FX) activation, which converts prothrombin to thrombin. During propagation, activated factor VIII (FVIII) activates FX in the presence of activated factor IX. Von Willebrand factor (vWF), produced by endothelial cells and megakaryocytes, has multiple roles in hemostasis. Its primary role is to serve as an adhesion molecule that anchors platelets to exposed collagen after endothelial cell damage. The factor also acts as a carrier protein of FVIII, thereby prolonging the half-life of FVIII.

Elevated circulating levels of FVIII and vWF are risk factors for venous thrombosis but the data supporting an association of FVII levels with arterial thrombosis are less consistent.<sup>1-5</sup> Hemorrhagic complications are associated with deficiency in FVII and vWF (von Willebrand disease), as well as X-linked deficiency in FVIII (Hemophilia A).<sup>6-9</sup> Plasma levels of these proteins are affected by environmental factors but they also are genetically influenced.<sup>10-13</sup> Heritability estimates range from 0.53-0.63 for FVII, 0.40-0.61 for FVIII, and 0.31-0.75 for vWF.<sup>12, 13</sup> To date, our understanding of genetic variation influencing plasma levels has been focused primarily on *cis*-acting variation in the genes encoding each protein product (*F7*, *F8*, and *VWF*, respectively). A large-scale genome-wide investigation of the genomic correlates of plasma levels has not been previously published. Using data from 23,608 adults, we investigated genome-wide associations between common genetic variation and plasma levels of FVII, FVIII, and vWF.

## Methods

### Setting

The discovery meta-analysis was conducted in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium, which includes data from 5 prospective, population-based cohorts of adults in the US and Europe.<sup>14</sup> Phenotype data were available from 4 of the cohorts: the Atherosclerosis Risk in Communities (ARIC) Study, the Cardiovascular Health Study (CHS), the Framingham Heart Study (FHS), and the Rotterdam Study (RS). In addition, participants from the British 1958 Birth Cohort (B58C) who had genome-wide data and were used as controls for the Wellcome Trust Case-Control Consortium were included in these analyses.<sup>15</sup> The designs of these studies have been described elsewhere.<sup>15-22</sup>

### Subjects

Eligible participants for these analyses had high-quality data from the genome-wide scans (see below), had at least 1 of the 3 phenotypes measured, and were not using a coumarin-based anticoagulant at the time of the phenotype measurement. Participants were of European (n=23,608) ancestry by self report. Each study received IRB approval and all participants provided written informed consent for the use of their DNA in research.

### Measures

**Coagulation Measures**—Plasma measures of FVII, FVIII, and vWF were obtained at the time of cohort entry for ARIC, CHS, and RS (except vWF, which was measured at visit 3), and at a follow-up visit for B58C (examination 2002 and 2003) and FHS (examination cycle 5, 1991-1995). Only the FVII phenotype in CHS was measured more than once and this included 2 measures, which were averaged for 2653 participants (81% of 3266 participants). The FVII phenotype was measured in 4 cohorts and included both antigen (FHS<sup>23</sup>) and activity (ARIC,<sup>24, 25</sup> CHS,<sup>26</sup> RS<sup>27</sup>). Factor VIII activity was measured in 3 cohorts (ARIC,<sup>24, 25</sup> CHS,<sup>26</sup> and RS) and vWF antigen was measured in 4 cohorts (ARIC,<sup>24, 25</sup>

B58C,<sup>28</sup> FHS,<sup>29</sup> and RS). Supplemental Table S1 provides details about assays in each of the 5 cohorts.

Baseline measures of clinical and demographic characteristics were obtained at the time of cohort entry for ARIC, CHS, and RS, and at the time of phenotype measurement for B58C and FHS. Measures, taken using standardized methods as specified by each study, included in-person measures of height and weight; and self-reported treatment of diabetes and hypertension, current alcohol consumption, and prevalent cardiovascular diseases (history of myocardial infarction, angina, coronary revascularization, stroke, or transient ischemic attack).

**Genotyping and Imputation**—Genotyping used DNA collected from phlebotomy (ARIC, CHS, FHS, RS) or cell lines (B58C). Genome-wide assays of SNPs were conducted independently in each cohort using various technologies: Affymetrix 6.0 for ARIC, Affymetrix 500K for B58C, Illumina 370CNV for CHS, Affymetrix 500K and gene-centric 50K for FHS, and Illumina 550K v3 for RS. Genotype quality control and data cleaning that included assessing Hardy-Weinberg equilibrium and variant call rates, were conducted independently by each study; details have been published elsewhere and are also provided in Supplemental Table S2.14

We investigated genetic variation in the 22 autosomal chromosomes and the X chromosome. Genotypes were coded as 0, 1, and 2 representing the number of copies of the coded alleles for all chromosomes except the X chromosome in males where genotypes were coded as 0 and 2.30 Each study independently imputed their genotype data to the ~2.6 million SNPs identified in the HapMap Caucasian (CEU) sample from the Centre d'Etude du Polymorphisme Humain.<sup>31-33</sup> Imputation software included MACH, BIMBAM or IMPUTE, with SNPs that pass quality-control criteria (Supplemental Table S2) were used among studies to impute unmeasured genotypes based on phased haplotypes observed in HapMap. Imputed data for the X chromosome were not available in ARIC or RS. Imputation results were summarized as an “allele dosage” defined as the expected number of copies of the minor allele at that SNP (a continuous value between 0 and 2) for each genotype. Each cohort calculated a ratio of observed to expected variance (OEV) of the dosage statistic for each SNP. This value, which generally ranges from 0 to 1 (poor to excellent), reflects imputation quality.

## Statistical Analyses

Investigators from all cohorts developed the pre-specified analytic plan described below. Each study independently analyzed their genotype-phenotype data. All studies used linear regression to conduct association analyses between measured and imputed SNPs and untransformed phenotype measures except for FHS, which used a linear mixed effects model to account for family relationships.<sup>34</sup> An additive genetic model with 1 degree of freedom was adjusted for age and sex. In addition, CHS and ARIC adjusted for field site and FHS adjusted for generation and ancestry using principal components.<sup>35</sup> For each analysis a genomic control coefficient, that estimated the extent of underlying population structure based on test-statistic inflation, was used to adjust standard errors.<sup>36</sup>

For each phenotype, within-study findings were combined across studies to produce summary results using standard meta-analytic approaches. For vWF antigen and FVIII activity levels, fixed-effects, inverse-variance weighted meta-analysis was performed, and summary p-values and  $\beta$ -coefficients were calculated. Parameter coefficient represents change (% vWF antigen or % FVIII activity) associated with 1-unit change in allele dosage. This approach was not appropriate for the FVII phenotype which combined analysis of activity and antigen measures. Instead effective-sample-size weighted meta-analysis was



performed to estimate summary p-values only. All meta-analyses were conducted using METAL software (<http://www.sph.umich.edu/csg/abecasis/METAL/index.html>). For loci containing genes already known to be associated with the phenotype, we conducted secondary analyses, adjusting for 1 or more SNPs within the known gene. This allowed us to assess possible novel associations in neighboring genes independently of the strong signal in the known SNPs.

The *a priori* threshold of genome-wide significance was set at a p-value of  $5.0 \times 10^{-8}$ . When more than 1 SNP clustered at a locus, we picked the SNP with the smallest p-value to represent the locus and to be used for replication. For each phenotype, the amount of variation explained by the top SNPs was the difference in the  $R^2$  value when comparing a model with adjustment variables only to a model that also included top genome-wide significant SNPs. Each study calculated the amount of variation explained and these estimates were combined across cohorts using weighted averages. Linkage disequilibrium (LD) between SNPs was estimated using a weighted average across cohorts.

## Replication

Novel genome-wide significant loci were subject to replication in 5 new populations. These populations included 1,375 ARIC participants not in the discovery cohort, 2,484 B58C cohort participants not in the discovery cohort, 1,006 participants from the Twins UK Study (587 from TUK-1 and 419 from TUK-2), 37 765 participants from VIS (Vis Croatia Study), 38 677 participants from ORCADES (Orkney Complex Disease Study),<sup>39</sup> and 1297 healthy, population-based Swedish subjects recruited to serve as controls in the PROCARDIS program.<sup>40</sup> The FVII replications were conducted using participants from ARIC, Twins UK, and PROCARDIS; for vWF, participants from ARIC, B58C, Twins UK, VIS, and ORCADES were included; and for FVIII, ARIC and Twins UK studies were included. Details about hemostasis assays can be found in Supplemental Table S1. Replication  $\beta$ -coefficients and p-values were meta-analyzed using samples-size (FVII) and inverse-variance (FVIII, vWF) weighting. The threshold of significance was  $5.0 \times 10^{-2}$ .

## Results

A total of 23,608 participants of European ancestry, were eligible for 1 or more analyses. Counts of participants and their characteristics are provided in Table 1. The average age ranged from 44.9 to 72.3 years for the 5 cohorts, and 54% of the participants were women. The median and interquartile range for each phenotype are also listed in Table 1.

## Factor VII Discovery

Within the 4 cohorts that conducted FVII analyses (ARIC, CHS, FHS, and RS), the genomic control coefficients for analyses were small ( $<1.04$ ), suggesting negligible test statistic inflation. Figure 1a presents all 2,734,954 meta-analysis p-values organized by chromosome and genomic position. Among these SNPs, 305 exceeded the genome-wide significance threshold and marked 5 regions on 5 chromosomes: 2p23, 4q25, 11q12, 13q34 (includes *F7*), and 20q11.2. A sixth region on chromosome 15 was marked by 6 SNPs all with very low MAFs ( $<0.007$ ) so this locus was not investigated further.

Table 2 lists the top SNP for each chromosomal region. Cohort-specific p-values are provided in Table 3. The amount of variance in the FVII phenotype explained by the 5 loci was 7.7%. Genome-wide significant signals at chromosomal position 2p23 (Supplemental Figure S1) were within or close to 1 gene: *GCKR* (glucokinase [hexokinase 4] regulator). *R*s1260326 was associated with the smallest p-value ( $6.2 \times 10^{-24}$ ; overall minor allele frequency [MAF] = 0.422) in this region, and this SNP codes a non-synonymous

substitution (proline to leucine) at amino acid position 446 in *GCKR*. Genome-wide significant signals at chromosomal position 4q25 (Supplemental Figure S2) were within the region of 2 genes: *ADH4* (alcohol dehydrogenase 4 [class II], pi polypeptide) and *ADH5* (alcohol dehydrogenase 5 [class III], chi polypeptide). *Rs1126670*, intronic to *ADH4*, had the smallest p-value ( $3.6 \times 10^{-12}$ ; MAF = 0.316). This SNP was in high LD ( $r^2=0.77$ ) with the top SNP in *ADH5* (*rs896992*; p-value =  $3.2 \times 10^{-11}$ ). At chromosomal position 11q12 (Supplemental Figure S3), genome-wide significant signals were within 2 genes: *MS4A2* and *MS4A6A* (membrane-spanning 4-domains, subfamily A, member 2 and member 6A, respectively). *Rs11230180* had the smallest p-value ( $2.0 \times 10^{-10}$ ) and was close to the *MS4A6A* gene and in high LD with top SNPs in *MS4A6A* ( $r^2=0.97$  for *rs17602572* [ $7.3 \times 10^{-10}$ ]) and in *MS4A2* ( $r^2=0.84$  for *rs2847666* [ $3.8 \times 10^{-8}$ ]).

The 63 genome-wide significant signals at chromosomal position 13q34 (Supplemental Figure S4) were within the factor VII structural gene (*F7*) and several other genes. *Rs488703* had the smallest p-value ( $9.0 \times 10^{-259}$ ) and is located in an *F7* intron. After adjusting for *rs6046* ( $4.4 \times 10^{-259}$ ), a missense SNP which leads to the arginine-glutamine (R353Q) functional substitution, 22 SNPs in the 13q34 region retained their genome-wide significance. These SNPs were within 2 genes: *rs3211727* ( $5.3 \times 10^{-23}$ ) in *F10* (coagulation factor X) and *rs1755693* ( $3.1 \times 10^{-19}$ ) in *MCF2L* (MCF.2 cell line derived transforming sequence-like).

The final chromosomal region containing genome-wide signals for FVII levels was position 20q11.2 (Supplemental Figure S5). *Rs867186* in *PROCR* had the smallest p-value ( $5.7 \times 10^{-37}$ ), and this variant leads to the serine-glycine (S219G) substitution in exon 4. The region also covered another 7 genes that contained SNPs with p-values that exceeded genome-wide significance: *ITCH* (itchy E3 ubiquitin protein ligase homolog), *PIGU* (phosphatidylinositol glycan anchor biosynthesis, class U), *ACSS2* (acyl-CoA synthetase short-chain family member 2), *MYH7B* (myosin, heavy chain 7B, cardiac muscle, beta), *TRPC4AP* (transient receptor potential cation channel, subfamily C, member 4 associated protein), *EDEM2* (ER degradation enhancer, mannosidase alpha-like 2), and *MMP24* (matrix metalloproteinase 24). As depicted on Supplemental Figure e, all top SNPs were in strong LD with *rs867186*.

Across all studies, significant heterogeneity of effect (p-value < 0.01) was detected for *rs1126670* and *rs488703*. Effect estimates were in the same direction but magnitudes differed between studies (Table 3a). When we restricted the discovery cohort to include only those studies that measured FVII activity (ARIC, CHS, and RS), p-values weakened for *rs1260326* ( $1.7 \times 10^{-18}$ ) and *rs867186* ( $2.8 \times 10^{-34}$ ), were virtually unchanged for *rs11230180* and *rs488703*, and strengthened for *rs1126670* ( $4.4 \times 10^{-14}$ ). A sixth locus emerged with 4 genome-wide significant SNPs at 11q13. Among these, *rs1149606* had the smallest p-value  $1.9 \times 10^{-9}$  (MAF = 0.176). This SNP was 4.0 kb from *TSKU* (tsukushin). This SNP had a p-value of  $2.2 \times 10^{-6}$  in the full discovery cohort. When we restricted the discovery cohort to the single study that measured measured FVII antigen, FHS, there were no SNPs that reached genome-wide significance.

## Factor VII Replication

The 4 new findings were tested in the ARIC, Twins UK (1 and 2), and Swedish PROCARDIS cohorts and all replicated. Replication p-values were  $6.5 \times 10^{-4}$  for *rs1260326*,  $3.8 \times 10^{-2}$  for *rs1126670*;  $1.4 \times 10^{-3}$  for *rs11230180*, and  $1.6 \times 10^{-7}$  for *rs867186* (see Supplemental Table S3a). When separating the FVII activity (ARIC, TUK-1) and FVII antigen (TUK-2 and PROCARDIS) cohorts, replication results were similar except *rs1260326* did not replicate in the activity cohorts (p-value  $9.0 \times 10^{-2}$ ).

## vWF and FVIII Discovery

Within each cohort that conducted genome-wide association analyses for vWF (ARIC, B58C, FHS, and RS) and FVIII (ARIC, CHS, and RS), the genomic control coefficients for analyses were small ( $<1.03$ ), suggesting negligible test statistic inflation. Figures 1b and 1c present all meta-analysis p-values (2,742,821 for vWF and 2,729,294 for FVIII) organized by chromosome and genomic position.

For vWF, 400 SNPs exceeded the genome-wide significance threshold and marked 8 regions on 6 chromosomes: 6q24, 8p21, 9q34, 12p13 (including *VWF*), 12q23, 12q24.3, 14q32 and 19p13.2. For FVIII, 191 SNPs exceeded the genome-wide significance threshold and marked 5 regions on 4 chromosomes: 6q24, 8p21, 9q34, 12p13, and 12q23. Table 2 lists the top SNP markers for each of the regions for both vWF and FVIII. Cohort-specific parameters are provided in Table 3. The 8 loci explained 12.8% of the vWF antigen variation and the 5 loci explained 10.0% of the FVIII activity variation. Because all the FVIII regions were a subset of those identified by vWF, we focus on the vWF findings.

Genome-wide significant signals at chromosomal position 6q24 (Supplemental Figure S6) where within or close to 1 gene: *STXBP5* (syntaxin binding protein 5). *Rs9390459*, with the smallest p-value ( $1.2 \times 10^{-22}$ ; MAF = 0.442) at this locus, encodes a synonymous substitution at amino acid position 779 in *STXBP5*. At chromosomal position 8p21 (Supplemental Figure S7), genome-wide significant signals were located on 1 gene: *SCARA5* (scavenger receptor class A, member 5). *Rs2726953* had the smallest p-value ( $1.3 \times 10^{-16}$ ; MAF = 0.309) and was intronic to *SCARA5*. Genome-wide significant signals at chromosomal position 9q34 (Supplemental Figure S8) were within 11 genes, 1 of which was *ABO*. All the SNPs with the very smallest p-values ( $<5.0 \times 10^{-324}$ ) were within *ABO*. After adjusting for 4 SNPs (*rs514659*, *rs8176749*, *rs8176704*, and *rs512770*) that collectively serve as tag-SNPs for the O, O<sup>v/2</sup>, A<sup>1-1/2</sup>, A<sup>2</sup>, and B haplotypes of *ABO*, 41 none of the 161 remaining SNPs at 9q34 was of genome-wide significance.

Three regions containing high-signal SNPs were identified on chromosome 12: 12p13, 12q23, and 12q24.3. The vWF structural gene resides within 12p13 (Supplemental Figure S9) and *rs1063857*, which leads to a synonymous amino acid substitution at position 795 of *VWF*, had the smallest p-value ( $1.7 \times 10^{-32}$ ; MAF = 0.360). Two genome-wide significant SNPs were located at 12q23 (Supplemental Figure S10) and were within 2 genes: *STAB2* (stabilin 2) and *NT5DC3* (5'-nucleotidase domain containing 3). *Rs4981022* in *STAB2* had the smaller p-value ( $7.3 \times 10^{-10}$ ; MAF = 0.315) and was intronic. The LD between this SNP and *rs10778286* ( $4.4 \times 10^{-8}$ ; MAF = 0.296) in *NT5DC3* was not complete ( $r^2=0.49$ ). At chromosomal region 12q24.3 (Supplemental Figure S11), genome-wide significant SNPs were found in 1 gene: *STX2* (syntaxin 2). *Rs7978987* had the smallest p-value ( $3.8 \times 10^{-11}$ ; MAF = 0.349) and was intronic to *STX2*. At chromosomal region 14q32 (Supplemental Figure S12), genome-wide significant SNPs were found within or close to 1 gene: *TC2N* (tandem C2 domains, nuclear). *Rs10133762* had the smallest p-value ( $2.3 \times 10^{-10}$ ; MAF = 0.448) and was intronic to *TC2N*. Lastly, genome-wide significant SNPs at chromosomal position 19p13.2 (Supplemental Figure S13) were located within 1 gene: *CLEC4M* (C-type lectin domain family 4, member M). *Rs868875* had the smallest p-value ( $1.3 \times 10^{-9}$ ; MAF = 0.262) and was intronic to *CLEC4M*.

The change in vWF associated with each additional copy of the minor allele was large for the ABO locus (increase of 24.1 in vWF antigen %) and was smaller for the other SNPs of genome-wide significance, ranging from 3.1 to 6.0 (Table 2). These change values were generally smaller for the FVIII phenotype compared with vWF. Only the ABO locus (*rs687621*) demonstrated heterogeneity of effect size across studies (Table 3b).



## Factor vWF Replication

The 6 new findings were tested in the ARIC, B58C, Twins UK (1 and 2), VIS, and ORCADES cohorts and 5 replicated. Replication p-values were  $1.7 \times 10^{-6}$  for rs9390459,  $3.1 \times 10^{-5}$  for rs2726953,  $8.7 \times 10^{-8}$  for rs4981022,  $1.9 \times 10^{-1}$  for rs7978987,  $3.3 \times 10^{-5}$  for rs10133762, and  $1.4 \times 10^{-2}$  for rs868875 (see Supplemental Table S3b).

## Discussion

In addition to confirming previously known associations of FVII levels with *F7* variation, and vWF and FVIII levels with *ABO* and *VWF* variation, respectively, this meta-analysis of data from 23,608 subjects of European ancestry identified novel, genome-wide significant associations of 4 loci with circulating FVII levels, 6 loci with vWF levels, and 3 loci with FVIII levels. The FVIII loci were a subset of the vWF loci. Several of these discoveries were associated with p-values many magnitudes smaller than the threshold of genome-wide significance and included new candidate genes for levels of FVII (*GCKR*, *ADH4*, *MS4A6A* and *PROCR*), of vWF (*STXBP5*, *SCARA5*, *STAB2*, *STX2*, *TC2N*, and *CLEC4M*), and of FVIII (*STXBP5*, *SCARA5*, *STAB2*). There was independent evidence for replication for 9 of 10 new findings in new populations.

### Genes Associated with Genome-wide Significant Findings: Factor VII

The strongest genetic associations with FVII levels resided in *F7* and confirm previously reported associations of SNPs in this locus.<sup>42, 43</sup> After adjustment for rs6046 (R353Q), there remained residual signal in this region with SNPs in *MCF2L*, *F7*, and *F10* reaching genome-wide significance. Upon further exploration, the residual signal was only detected in 1 of the 4 cohorts (ARIC), where the rs6046 SNP was imputed with only modest precision (OEV = 0.607). Although the residual signal is likely attributable to the incomplete adjustments, we cannot rule out the possibility that other sequence variations in the 13q34 region contribute to FVII levels.<sup>44</sup>

The association of variants in *PROCR* with FVII levels has not been reported previously and is novel. Recent reports have described activated FVII (FVIIa) binding to the endothelial protein C receptor via the Gla domain, which has homologous counterpart on the protein C molecule.<sup>45, 46</sup> This binding appears to inhibit procoagulant activity of the FVIIa-tissue factor complex and may impact the clearance of FVIIa and receptor signaling through competition with the binding site. Variation in *GCKR*, whose gene product inhibits glucokinase in liver and pancreatic islet cells, has not been associated previously with FVII levels but has been associated with C-reactive protein and triglyceride levels.<sup>47, 48</sup> Factor VII protein concentrations are associated with triglyceride levels in the fasting state, but there is a direct effect of plasma triglycerides on FVII activity.<sup>49</sup> Postprandial increase in plasma triglycerides is directly correlated with increase in FVII activity, but the protein concentration remains constant. Genetic variation in alcohol dehydrogenase 3 has been associated with myocardial infarction risk and coronary heart disease risk.<sup>50, 51</sup> Plasma levels of FVII are negatively correlated with alcohol consumption and *ADH4*, a gene that regulates alcohol metabolism, and may have an effect on FVII levels. There are no previous reports of associations with levels of FVII of genetic variation in *ADH4*, *ADH5* or other alcohol dehydrogenase genes. Little has been published on the *MS4A6A* gene.<sup>52</sup>

When comparing findings by the different measurements of FVII, either antigen or activity level, we found little evidence that the genetic predictors we identified differed by sub-phenotype. Nonetheless a more extensive comparison of the 2 sub-phenotypes is merited and would require a larger representation of FVII antigen measures in the discovery cohort.

## Genes Associated with Genome-wide Significant Findings: Factor vWF and FVIII

Because the vWF molecule transports FVIII and the 2 factors are highly correlated, we expected overlap in some of the vWF and FVIII findings.<sup>11</sup> No unique genetic predictors of FVIII levels on the 22 autosomal chromosomes or the X chromosome, where *F8* is located, were identified. This may indicate that in healthy individuals genetic determinants of FVIII levels are primarily dependant upon vWF levels.

The strong associations between SNPs located in *ABO* and *VWF* and FVIII levels were expected since it is known that individuals with blood group O have 25-30% lower vWF levels than individuals with non-O blood groups.<sup>53</sup> The 2 SNPs in *VWF* with the smallest p-values for FVIII (*rs1063856*) and vWF (*rs1063857*) levels are in strong LD and mark non-synonymous and synonymous substitutions at amino acids 789 and 795, respectively. These variations are located in exon 18, which encodes for the D' domain, and are involved in binding of FVIII. Rare genetic variation in *VWF* has been associated low levels of vWF, which characterize von Willebrand disease, either through lowering protein expression or through increased clearance. Rare missense mutations in the D' part of the vWF gene have been associated with normal or reduced levels of vWF and low levels of FVIII, which is characteristic of von Willebrand disease type 2N.<sup>54, 55</sup> Rare genetic variation in the *VWF* has also been associated with low levels of vWF, which characterize von Willebrand disease, either through lowering the proteins expression or through increased clearance.

We detected a novel and relatively strong association for the *STXBP5* gene and also for *STX2*, the binding substrate for *STXBP5*, with vWF and FVIII levels. We were not able to replicate the STX2 finding. Syntaxin 4, one of the soluble NSF attachment protein receptor (SNARE) proteins, has been linked to exocytosis of Weibel Palade bodies (WPB) in endothelial cells by targeting WPB to the plasma membrane.<sup>56</sup> Knockdown of syntaxin 4 resulted in inhibition of WPB exocytosis.<sup>57</sup> Von Willebrand factor is the main constituent of WPB and upon stimulation, for instance due to stress, or after endothelial damage WPB are excreted by endothelial cells, and give rise to increased plasma vWF levels. Syntaxin 2 and vertebrate *STXBP5* have not been shown to be involved in vWF secretion, however our observations may indicate a functional role of these proteins in determining circulating vWF levels.

Scavenger receptor Class A member 5 (*SCARA5*) belongs to the group of scavenger molecules, that have as primary function the initiation of immune responses. *SCARA5* has recently been characterized in more detail and is solely expressed in epithelial cells. It has not yet been linked to vWF, and the nature of the genetic associations found in our study is yet unknown. Stabilin-2 (*STAB2*) is a transmembrane receptor protein primarily expressed in liver and spleen sinusoidal endothelial cells. The receptor binds and endocytoses various ligands including heparin, low density lipoproteins, bacteria and advanced glycosylation products.<sup>58</sup> Taken together, these results generate the hypothesis that several loci may be involved in vWF and FVIII clearance or uptake.

## Relevance for Cardiovascular Outcomes

The relevance of the reported associations of genome-wide variation with levels of hemostasis proteins to cardiovascular endpoints will require expanded research efforts. Elevated circulating levels of FVIII and vWF are independent risk factors for venous thrombosis and the genetic variants identified here explained roughly 10% of the variance in these measures.<sup>4</sup> Variation in *ABO* is a known risk factor for venous thrombosis and the risk associated with the newly identified candidate genes presented here can only be hypothesized.<sup>59, 60</sup> The association between FVII levels and thrombotic risk is less clear and epidemiologic findings have not been consistent.<sup>1-3, 5</sup> Similarly, published evidence

supporting the role of *F7* variation in atherothrombotic risk has been heterogeneous and an exploration of newly identified candidate genes in this report has not been undertaken.<sup>61, 62</sup>

### Strengths and Limitations

This is the first genome-wide association study to attempt to discover novel genetic associations with the 3 hemostasis phenotypes: FVII, FVIII, and vWF. It included 23,608 individuals of European ancestry and ~2.6 million markers spread throughout the genome. Data came from 5 large, population-based cohort studies where cardiovascular outcomes were primary study endpoints. Phenotypes were measured in a standardized fashion within each cohort but measures between cohorts differed and likely introduced between-group variability. This variability may decrease statistical power to find associations of smaller magnitudes.

Supplemental Table S4 lists the minimum percent change in the hemostasis measures for variants with a range of MAFs and assuming power of 80% or greater and an  $\alpha$  of  $5.0 \times 10^{-8}$ . For a SNP with an MAF of 0.05, we had sufficient power to detect a 5% change in FVII, a 6% change in FVIII, and a 7% change in vWF. Not all SNPs tested were directly genotyped and the imputation quality varied across SNPs. For poorly imputed SNPs, there was reduced statistical power to detect an association. For each identified locus, we chose the SNP with the smallest p-value but the causal variant—if one exists—need not be the one with the smallest p-value and effect-sizes associated with the SNP may be an overestimate of any true effect. In situations where genes clustered near a gene containing 1 or more variants that were strongly associated with the phenotype, we used statistical adjustment to determine if neighboring genes had independent associations with the phenotype. This approach has limitations and cannot identify novel causal variants that are in LD with known causal variants.

### Summary

Using data from 5 community-based cohorts that included 23,608 participants, we identified 4 novel genetic associations with FVII activity and antigen levels and 6 novel genetic associations with vWF antigen levels: 9 of the 10 findings were replicated. New candidate genes of interest were discovered outside known biologic pathways influencing these essential hemostasis factors, and our findings may point to new pathways to target for the prevention and treatment of hemorrhagic and thrombotic disorders.

#### Commentary

Elevated circulating levels of coagulation factor VIII (FVIII) and von Willebrand factor (vWF) are risk factors for venous thrombosis but the data supporting an association of coagulation factor VII (FVII) levels with arterial thrombosis are less consistent. Hemorrhagic complications are associated with deficiency in FVII and vWF (von Willebrand disease), as well as X-linked deficiency in FVIII (Hemophilia A). To date, our understanding of genetic variation influencing plasma levels has been focused primarily on variation in the genes encoding each protein product (*F7*, *F8*, and *VWF*, respectively). Using data from 23,608 adults drawn from community populations, we investigated genome-wide associations between common genetic variation and plasma levels of FVII, FVIII, and vWF. For FVII, we identified 5 loci on 5 chromosomes that exceeded genome-wide significance. All loci were within or near genes, including 4 new candidate genes and *F7*. For vWF, we identified 8 loci on 6 chromosomes. All loci were within genes, including 6 new candidate genes, as well as *ABO* and *VWF*. For FVIII, 5 loci were identified and overlapped vWF findings. We replicated 9 of the 10 new findings in independent samples. In summary, new genetic associations were discovered

in biologic pathways not previously associated with circulating levels of these factors, including proteins implicated in uptake and intracellular transport of the factors. These findings may point to novel prevention and treatment targets of hemostasis disorders.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors acknowledge the essential role of the CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium in development and support of this manuscript. CHARGE members include National Heart, Lung, and Blood Institute's (NHLBI) Atherosclerosis Risk in Communities (ARIC) Study, NIA's Iceland Age, Gene/Environment Susceptibility (AGES) Study, NHLBI's Cardiovascular Health Study and Framingham Heart Study, and the Netherlands Rotterdam Study. The authors also acknowledge the thousands of study participants who volunteered their time to help advance science and the scores of research staff and scientists who have made this research possible. Further, ARIC would like to thank the University of Minnesota Supercomputing Institute for use of the blade supercomputers; Twins UK would like to acknowledge essential contribution of Peter Grant and Angela Carter from the Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, UK for measurements of clotting factors phenotypes; and VIS would like to acknowledge the invaluable contributions of the recruitment team (including those from the Institute of Anthropological Research in Zagreb) in Vis, the administrative teams in Croatia and Edinburgh and the people of Vis. Genotyping was performed at the Wellcome Trust Clinical Research Facility in Edinburgh.

### Funding Sources

The Atherosclerosis Risk in Communities Study is supported by National Heart, Lung, and Blood Institute (NHLBI) contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, and N01-HC-55022 and grants R01-HL-087641, R01-HL-59367 and R01-HL-086694; National Human Genome Research Institute contract U01-HG-004402; and NIH contract HHSN268200625226C. The infrastructure was partly supported by grant number UL1-RR-025005, a component of the NIH and NIH Roadmap for Medical Research.

We acknowledge use of phenotype and genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. (<http://www.b58cgene.sgu.ac.uk/>).

The Cardiovascular Health Study (CHS) is supported by contract numbers N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, grant numbers U01 HL080295 and R01 HL 087652 from NHLBI, with additional contribution from the National Institute of Neurological Disorders and Stroke. A full list of principal CHS investigators and institutions can be found at <http://www.chs-nhlbi.org/pi.htm>. Support was also provided by NHLBI grants HL073410 and HL095080 and the Leducq Foundation, Paris, France for the development of Transatlantic Networks of Excellence in Cardiovascular Research. DNA handling and genotyping was supported in part by National Center for Research Resources grant M01-RR00425 to the Cedars-Sinai General Clinical Research Center Genotyping core and National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

This research was conducted in part using data and resources from the Framingham Heart Study of the NHLBI of the NIH and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. Partial investigator support was provided by National Institute of Diabetes and Digestive and Kidney Diseases K24 DK080140 (JB Meigs). This work was partially supported by NHLBI's Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc., for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center.

The Rotterdam Study is supported by the Erasmus Medical Center and Erasmus University Rotterdam; the Netherlands Organization for Scientific Research; the Netherlands Organization for Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly; The Netherlands Heart Foundation; the Ministry of Education, Culture and Science; the Ministry of Health Welfare and Sports; the European Commission; and the Municipality of Rotterdam. Support for genotyping was provided by the Netherlands Organization for

Scientific Research (NWO) (175.010.2005.011, 911.03.012) and Research Institute for Diseases in the Elderly (RIDE). This study was further supported by the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. Dr. Dehghan is supported by NWO, RIDE (94800022).

The Twins UK study was funded by the Wellcome Trust; European Community's Sixth and Seventh Framework Programmes (FP-6/2005-2008) LIFE SCIENCES & HEALTH (Ref 005268 Genetic regulation of the end stage clotting process that leads to thrombotic stroke: The EuroClot Consortium and (FP7/2007-2013), ENGAGE project HEALTH-F4-2007-201413 and the FP-5 GenomeUtwinn Project (QLG2-CT-2002-01254). The study also receives support from the Dept of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London. TDS is an NIHR Senior Investigator. The project also received support from a Biotechnology and Biological Sciences Research Council (BBSRC) project grant. (G20234). The authors acknowledge the funding and support of the National Eye Institute via an NIH/CIDR genotyping project (PI: Terri Young).

The VIS study in the Croatian island of Vis was supported through the grants from the Medical Research Council UK and Ministry of Science, Education and Sport of the Republic of Croatia (number 108-1080315-0302) and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947).

ORCADES was supported by the Chief Scientist Office of the Scottish Government, the Royal Society and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh.

The PROCARDIS program was funded by the EC Sixth Framework Programme (LSHM-CT-2007-037273), the Swedish Research Council (8691), the Knut and Alice Wallenberg Foundation, the Swedish Heart-Lung Foundation, the Leducq Foundation, Paris, the Stockholm County Council (560283) and AstraZeneca. Genotyping of Swedish PROCARDIS control samples was performed at the SNP Technology Platform (head: Prof. Ann-Christine Syvänen), Department of Medical Sciences, Uppsala University, Sweden.

## References

1. Folsom AR. Hemostatic risk factors for atherothrombotic disease: an epidemiologic view. *Thromb Haemost.* 2001; 86:366–373. [PubMed: 11487026]
2. Smith A, Patterson C, Yarnell J, Rumley A, Ben-Shlomo Y, Lowe G. Which hemostatic markers add to the predictive value of conventional risk factors for coronary heart disease and ischemic stroke? The Caerphilly Study. *Circulation.* 2005; 112:3080–3087. [PubMed: 16286603]
3. Folsom AR, Cushman M, Heckbert SR, Ohira T, Rasmussen-Torvik L, Tsai MY. Factor VII coagulant activity, factor VII -670A/C and -402G/A polymorphisms, and risk of venous thromboembolism. *J Thromb Haemost.* 2007; 5:1674–1678. [PubMed: 17663738]
4. Koster T, Blann AD, Briet E, Vandenbroucke JP, Rosendaal FR. Role of clotting factor VIII in effect of von Willebrand factor on occurrence of deep-vein thrombosis. *Lancet.* 1995; 345:152–155. [PubMed: 7823669]
5. Spiel AO, Gilbert JC, Jilma B. von Willebrand factor in cardiovascular disease: focus on acute coronary syndromes. *Circulation.* 2008; 117:1449–1459. [PubMed: 18347221]
6. Cooper DN, Millar DS, Wacey A, Banner DW, Tuddenham EG. Inherited factor VII deficiency: molecular genetics and pathophysiology. *Thromb Haemost.* 1997; 78:151–160. [PubMed: 9198146]
7. Acharya SS, Coughlin A, Dimichele DM. Rare Bleeding Disorder Registry: deficiencies of factors II, V, VII, X, XIII, fibrinogen and dysfibrinogenemias. *J Thromb Haemost.* 2004; 2:248–256. [PubMed: 14995986]
8. Nichols WC, Ginsburg D. von Willebrand disease. *Medicine (Baltimore).* 1997; 76:1–20. [PubMed: 9064484]
9. Antonarakis SE, Youssoufian H, Kazazian HH Jr. Molecular genetics of hemophilia A in man (factor VIII deficiency). *Mol Biol Med.* 1987; 4:81–94. [PubMed: 2888005]
10. Folsom AR, Wu KK, Davis CE, Conlan MG, Sorlie PD, Szklo M. Population correlates of plasma fibrinogen and factor VII, putative cardiovascular risk factors. *Atherosclerosis.* 1991; 91:191–205. [PubMed: 1789804]
11. Conlan MG, Folsom AR, Finch A, Davis CE, Sorlie P, Marcucci G, Wu KK. Associations of factor VIII and von Willebrand factor with age, race, sex, and risk factors for atherosclerosis. The Atherosclerosis Risk in Communities (ARIC) Study. *Thromb Haemost.* 1993; 70:380–385. [PubMed: 8259533]

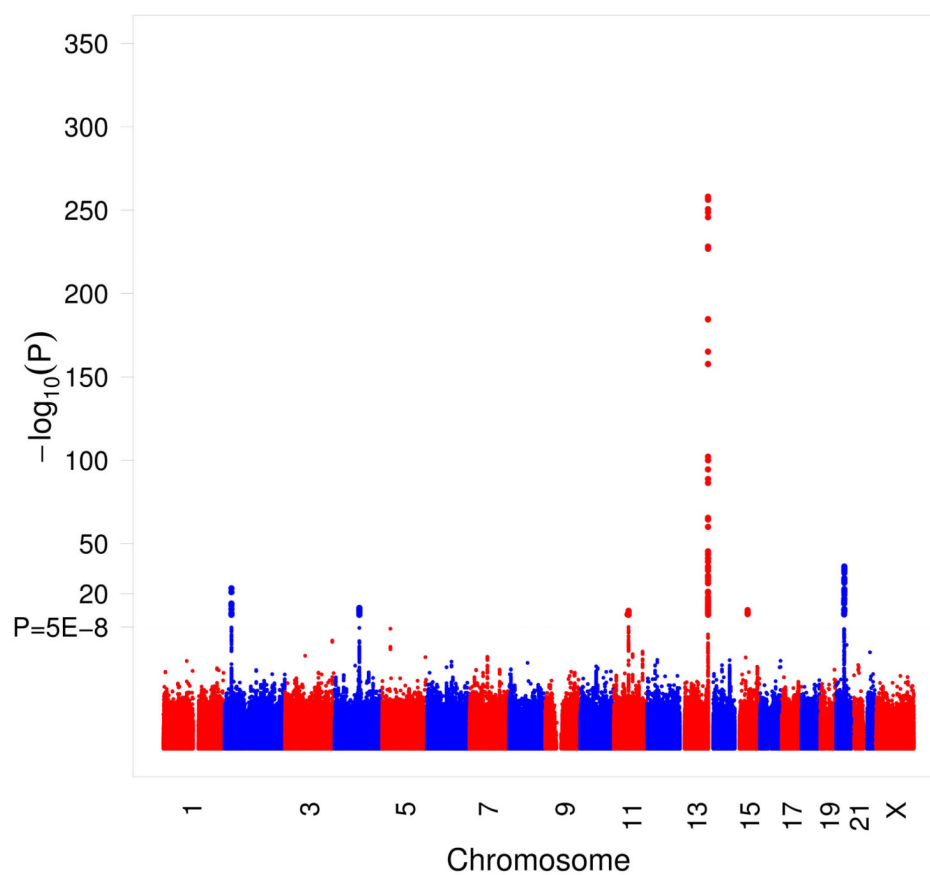


12. Souto JC, Almasy L, Borrell M, Gari M, Martinez E, Mateo J, Stone WH, Blangero J, Fontcuberta J. Genetic determinants of hemostasis phenotypes in Spanish families. *Circulation*. 2000; 101:1546–1551. [PubMed: 10747348]
13. de Lange M, Snieder H, Ariens RA, Spector TD, Grant PJ. The genetics of haemostasis: a twin study. *Lancet*. 2001; 357:101–105. [PubMed: 11197396]
14. Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JJ, Uitterlinden AG, Harris TB, Witteman JC, Boerwinkle E. Cohorts for Heart and Aging Research in Genomic Epidemiology: CHARGE Consortium. Design of prospective meta-analysis of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet*. 2009; 2:73–80. [PubMed: 20031568]
15. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007; 447:661–678. [PubMed: 17554300]
16. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol*. 1989; 129:687–702. [PubMed: 2646917]
17. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, Kuller LH, Manolio TA, Mittelmark MB, Newman A, O'Leary DH, Psaty BM, Rautaharju P, Tracy RP, Weiler PG. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol*. 1991; 1:263–276. [PubMed: 1669507]
18. Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health*. 1951; 41:279–281. [PubMed: 14819398]
19. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med*. 1975; 4:518–525. [PubMed: 1208363]
20. Hofman A, Breteler MM, van Duijn CM, Krestin GP, Pols HA, Stricker BH, Tiemeier H, Uitterlinden AG, Vingerling JR, Witteman JC. The Rotterdam Study: objectives and design update. *Eur J Epidemiol*. 2007; 22:819–829. [PubMed: 17955331]
21. Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol*. 1991; 7:403–422. [PubMed: 1833235]
22. Strachan DP, Rudnicka AR, Power C, Shepherd P, Fuller E, Davis A, Gibb I, Kumari M, Rumley A, Macfarlane GJ, Rahi J, Rodgers B, Stansfeld S. Lifecourse influences on health among British adults: effects of region of residence in childhood and adulthood. *Int J Epidemiol*. 2007; 36:522–531. [PubMed: 17255346]
23. Feng D, Tofler GH, Larson MG, O'Donnell CJ, Lipinska I, Schmitz C, Sutherland PA, Johnstone MT, Muller JE, D'Agostino RB, Levy D, Lindpaintner K. Factor VII gene polymorphism, factor VII levels, and prevalent cardiovascular disease: the Framingham Heart Study. *Arterioscler Thromb Vasc Biol*. 2000; 20:593–600. [PubMed: 10669660]
24. Folsom AR, Wu KK, Rosamond WD, Sharrett AR, Chambless LE. Prospective study of hemostatic factors and incidence of coronary heart disease: the Atherosclerosis Risk in Communities (ARIC) Study. *Circulation*. 1997; 96:1102–1108. [PubMed: 9286936]
25. Chambless LE, McMahon R, Wu K, Folsom A, Finch A, Shen YL. Short-term intraindividual variability in hemostasis factors. The ARIC Study. Atherosclerosis Risk in Communities Intraindividual Variability Study. *Ann Epidemiol*. 1992; 2:723–733. [PubMed: 1342324]
26. Cushman M, Yanez D, Psaty BM, Fried LP, Heiss G, Lee M, Polak JF, Savage PJ, Tracy RP. Association of fibrinogen and coagulation factors VII and VIII with cardiovascular risk factors in the elderly: the Cardiovascular Health Study. Cardiovascular Health Study Investigators. *Am J Epidemiol*. 1996; 143:665–676. [PubMed: 8651228]
27. Mennen LI, Witteman JC, den Breeijen JH, Schouten EG, de Jong PT, Hofman A, Grobbee DE. The association of dietary fat and fiber with coagulation factor VII in the elderly: the Rotterdam Study. *Am J Clin Nutr*. 1997; 65:732–736. [PubMed: 9062522]
28. Rudnicka AR, Rumley A, Lowe GD, Strachan DP. Diurnal, seasonal, and blood-processing patterns in levels of circulating fibrinogen, fibrin D-dimer, C-reactive protein, tissue plasminogen activator, and von Willebrand factor in a 45-year-old population. *Circulation*. 2007; 115:996–1003. [PubMed: 17296859]

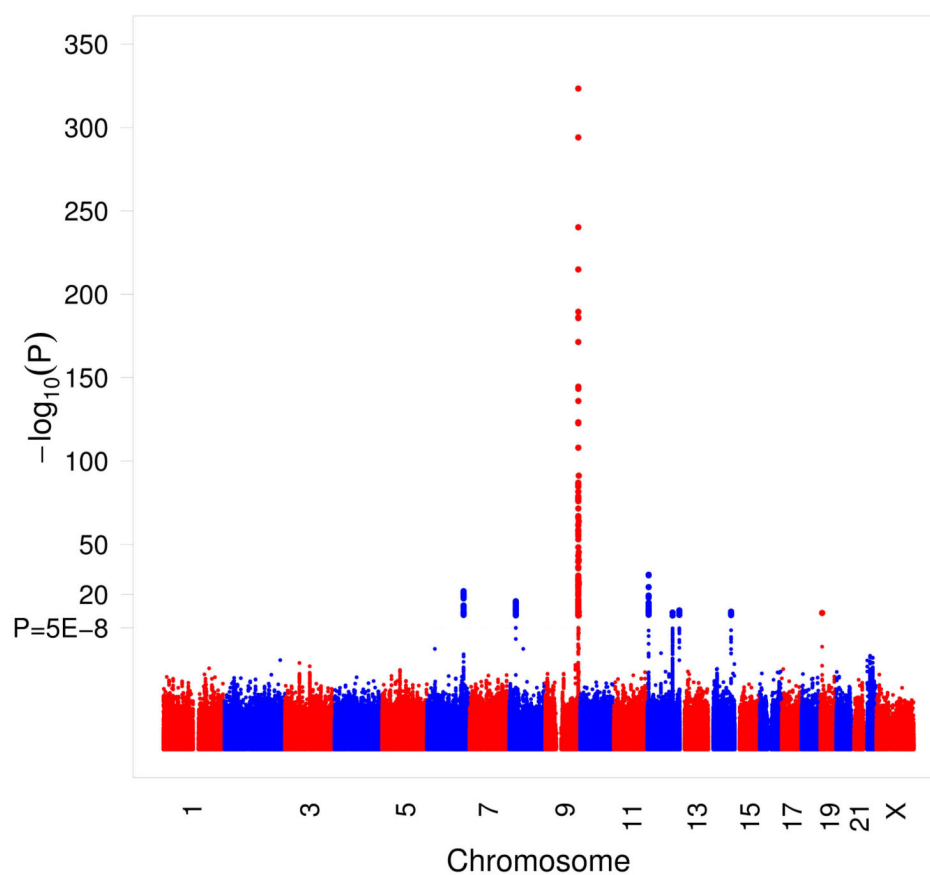
29. Tofler GH, Massaro J, Levy D, Mittleman M, Sutherland P, Lipinska I, Muller JE, D'Agostino RB. Relation of the prothrombotic state to increasing age (from the Framingham Offspring Study). *Am J Cardiol*. 2005; 96:1280–1283. [PubMed: 16253598]
30. Clayton D. Testing for association on the X chromosome. *Biostatistics*. 2008; 9:593–600. [PubMed: 18441336]
31. Servin B, Stephens M. Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genet*. 2007; 3:e114. [PubMed: 17676998]
32. Li Y, Abecasis GR. Mach 1.0: rapid haplotype reconstruction and missing genotype inference. *Am J Hum Genet*. 2006; 79:2290.
33. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*. 2007; 39:906–913. [PubMed: 17572673]
34. Abecasis GR, Cardon LR, Cookson WO, Sham PC, Cherny SS. Association analysis in a variance components framework. *Genet Epidemiol*. 2001; 21(Suppl 1):S341–346. [PubMed: 11793695]
35. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006; 38:904–909. [PubMed: 16862161]
36. Devlin B, Roeder K, Wasserman L. Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol*. 2001; 60:155–166. [PubMed: 11855950]
37. Andrew T, Hart DJ, Snieder H, de Lange M, Spector TD, MacGregor AJ. Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin Res*. 2001; 4:464–477. [PubMed: 11780939]
38. Vitart V, Rudan I, Hayward C, Gray NK, Floyd J, Palmer CN, Knott SA, Kolcic I, Polasek O, Graessler J, Wilson JF, Marinaki A, Riches PL, Shu X, Janicijevic B, Smolej-Narancic N, Gorgoni B, Morgan J, Campbell S, Biloglav Z, Barac-Lauc L, Pericic M, Klaric IM, Zgaga L, Skaric-Juric T, Wild SH, Richardson WA, Hohenstein P, Kimber CH, Tenesa A, Donnelly LA, Fairbanks LD, Aringer M, McKeigue PM, Ralston SH, Morris AD, Rudan P, Hastie ND, Campbell H, Wright AF. SLC2A9 is a newly identified urate transporter influencing serum urate concentration, urate excretion and gout. *Nat Genet*. 2008; 40:437–442. [PubMed: 18327257]
39. McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, Smolej-Narancic N, Janicijevic B, Polasek O, Tenesa A, Macleod AK, Farrington SM, Rudan P, Hayward C, Vitart V, Rudan I, Wild SH, Dunlop MG, Wright AF, Campbell H, Wilson JF. Runs of homozygosity in European populations. *Am J Hum Genet*. 2008; 83:359–372. [PubMed: 18760389]
40. Broadbent HM, Peden JF, Lorkowski S, Goel A, Ongen H, Green F, Clarke R, Collins R, Franzosi MG, Tognoni G, Seedorf U, Rust S, Eriksson P, Hamsten A, Farrall M, Watkins H. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. *Hum Mol Genet*. 2008; 17:806–814. [PubMed: 18048406]
41. Olsson ML, Chester MA. Polymorphism and recombination events at the ABO locus: a major challenge for genomic ABO blood grouping strategies. *Transfus Med*. 2001; 11:295–313. [PubMed: 11532186]
42. Humphries S, Temple A, Lane A, Green F, Cooper J, Miller G. Low plasma levels of factor VIIc and antigen are more strongly associated with the 10 base pair promoter (-323) insertion than the glutamine 353 variant. *Thromb Haemost*. 1996; 75:567–572. [PubMed: 8743180]
43. Kathiresan S, Yang Q, Larson MG, Camargo AL, Tofler GH, Hirschhorn JN, Gabriel SB, O'Donnell CJ. Common genetic variation in five thrombosis genes and relations to plasma hemostatic protein level and cardiovascular disease risk. *Arterioscler Thromb Vasc Biol*. 2006; 26:1405–1412. [PubMed: 16614319]
44. Vossen CY, Hasstedt SJ, Demers C, Rosendaal FR, Bovill EG. Linkage analysis for three coagulation factors clustering on chromosome 13q34: factor VII, factor X and protein Z. *J Thromb Haemost*. 2007; 5:1325–1327. [PubMed: 17403098]
45. Ghosh S, Pendurthi UR, Steinoe A, Esmon CT, Rao LV. Endothelial cell protein C receptor acts as a cellular receptor for factor VIIa on endothelium. *J Biol Chem*. 2007; 282:11849–11857. [PubMed: 17327234]

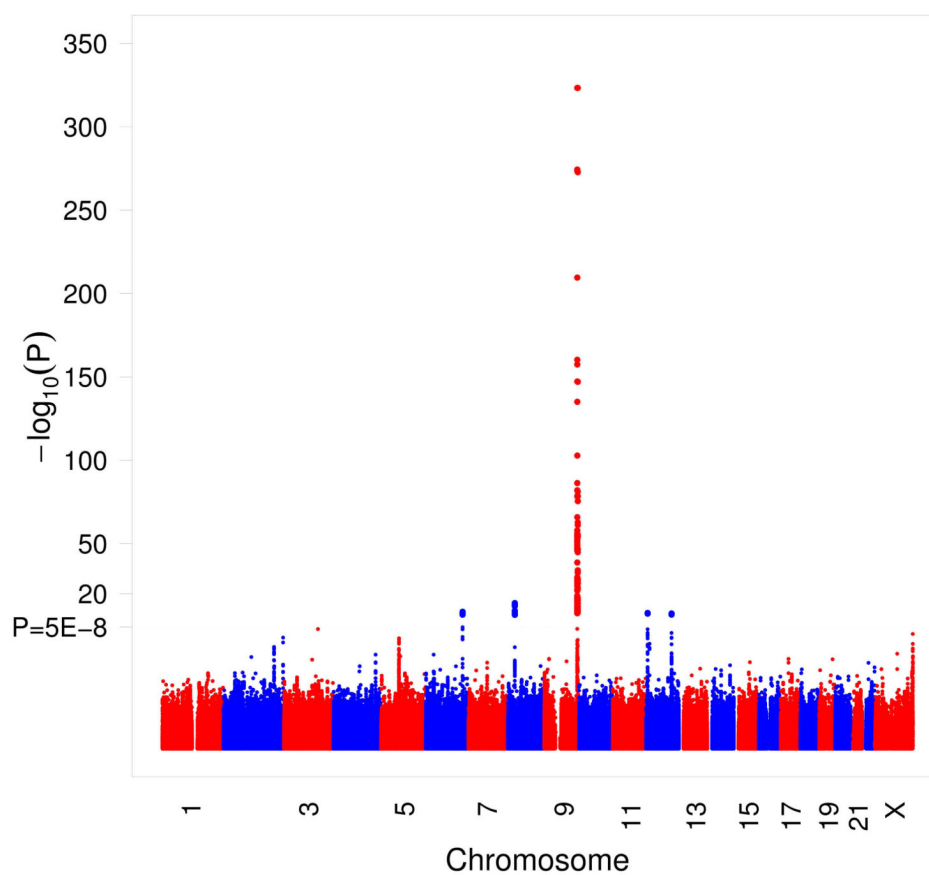
46. Lopez-Sagaseta J, Montes R, Puy C, Diez N, Fukudome K, Hermida J. Binding of factor VIIa to the endothelial cell protein C receptor reduces its coagulant activity. *J Thromb Haemost.* 2007; 5:1817–1824. [PubMed: 17723119]
47. Ridker PM, Pare G, Parker A, Zee RY, Danik JS, Buring JE, Kwiatkowski D, Cook NR, Miletich JP, Chasman DI. Loci related to metabolic-syndrome pathways including LEPR, HNF1A, IL6R, and GCKR associate with plasma C-reactive protein: the Women's Genome Health Study. *Am J Hum Genet.* 2008; 82:1185–1192. [PubMed: 18439548]
48. Saxena R, Voight BF, Lyssenko V, Burt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Riche D, Purcell S. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science.* 2007; 316:1331–1336. [PubMed: 17463246]
49. Skartlien AH, Lyberg-Beckmann S, Holme I, Hjermann I, Prydz H. Effect of alteration in triglyceride levels on factor VII-phospholipid complexes in plasma. *Arteriosclerosis.* 1989; 9:798–801. [PubMed: 2686604]
50. Hines LM, Stampfer MJ, Ma J, Gaziano JM, Ridker PM, Hankinson SE, Sacks F, Rimm EB, Hunter DJ. Genetic variation in alcohol dehydrogenase and the beneficial effect of moderate alcohol consumption on myocardial infarction. *N Engl J Med.* 2001; 344:549–555. [PubMed: 11207350]
51. Younis J, Cooper JA, Miller GJ, Humphries SE, Talmud PJ. Genetic variation in alcohol dehydrogenase 1C and the beneficial effect of alcohol intake on coronary heart disease risk in the Second Northwick Park Heart Study. *Atherosclerosis.* 2005; 180:225–232. [PubMed: 15910847]
52. Liang Y, Buckley TR, Tu L, Langdon SD, Tedder TF. Structural organization of the human MS4A gene cluster on Chromosome 11q12. *Immunogenetics.* 2001; 53:357–368. [PubMed: 11486273]
53. Gill JC, Endres-Brooks J, Bauer PJ, Marks WJ Jr. Montgomery RR. The effect of ABO blood group on the diagnosis of von Willebrand disease. *Blood.* 1987; 69:1691–1695. [PubMed: 3495304]
54. Mazurier C, Goudemand J, Hilbert L, Caron C, Fressinaud E, Meyer D. Type 2N von Willebrand disease: clinical manifestations, pathophysiology, laboratory diagnosis and molecular biology. *Best Pract Res Clin Haematol.* 2001; 14:337–347. [PubMed: 11686103]
55. James PD, Notley C, Hegadorn C, Leggo J, Tuttle A, Tinlin S, Brown C, Andrews C, Labelle A, Chirninian Y, O'Brien L, Othman M, Rivard G, Rapson D, Hough C, Lillicrap D. The mutational spectrum of type 1 von Willebrand disease: Results from a Canadian cohort study. *Blood.* 2007; 109:145–154. [PubMed: 17190853]
56. Lowenstein CJ, Morrell CN, Yamakuchi M. Regulation of Weibel-Palade body exocytosis. *Trends Cardiovasc Med.* 2005; 15:302–308. [PubMed: 16297768]
57. Fu J, Naren AP, Gao X, Ahmmed GU, Malik AB. Protease-activated receptor-1 activation of endothelial cells induces protein kinase C $\alpha$ -dependent phosphorylation of syntaxin 4 and Munc18c: role in signaling p-selectin expression. *J Biol Chem.* 2005; 280:3178–3184. [PubMed: 15576373]
58. Harris EN, Weigel PH. The ligand-binding profile of HARE: hyaluronan and chondroitin sulfates A, C, and D bind to overlapping sites distinct from the sites for heparin, acetylated low-density lipoprotein, dermatan sulfate, and CS-E. *Glycobiology.* 2008; 18:638–648. [PubMed: 18499864]
59. Morelli VM, De Visser MC, Vos HL, Bertina RM, Rosendaal FR. ABO blood group genotypes and the risk of venous thrombosis: effect of factor V Leiden. *J Thromb Haemost.* 2005; 3:183–185. [PubMed: 15634288]
60. Wiggins KL, Smith NL, Glazer NL, Rosendaal FR, Heckbert SR, Psaty BM, Rice KM, Lumley T. ABO genotype and risk of thrombotic events and hemorrhagic stroke. *J Thromb Haemost.* 2009; 7:263–269. [PubMed: 19036074]

61. Girelli D, Russo C, Ferraresi P, Olivieri O, Pinotti M, Friso S, Manzato F, Mazzucco A, Bernardi F, Corrocher R. Polymorphisms in the factor VII gene and the risk of myocardial infarction in patients with coronary artery disease. *N Engl J Med.* 2000; 343:774–780. [PubMed: 10984565]
62. Ye Z, Liu EH, Higgins JP, Keavney BD, Lowe GD, Collins R, Danesh J. Seven haemostatic gene polymorphisms in coronary disease: meta-analysis of 66,155 cases and 91,307 controls. *Lancet.* 2006; 367:651–658. [PubMed: 16503463]









**Figures 1a-c.**

The genome-wide  $\log_{10}$  p-value plots of FVII, FVIII, and vWF. The horizontal line marks the  $5.0 \times 10^{-8}$  p-value threshold of genome-wide significance.

**Table 1**  
Characteristics of the study participants at the time of hemostasis factor measurements

Characteristic	ARIC	B58C	CHS	FHS	RS	
					Original (RS I)	Extended (RS II)
Count	8053	1461	3272	2953	5974	1895
Age, yrs (SD)	54.3 (5.7)	44.9 (0.4)	72.3 (5.4)	54.5 (9.8)	69.4 (9.1)	64.8 (8.0)
Male, %	47.1	50.0	39.2	42.7	40.6	54.5
European ancestry participants, %	100	100	100	100	100	100
Body mass index, kg/m <sup>2</sup> (SD)	27.0 (4.9)	27.4 (4.8)	26.3 (4.4)	27.4 (5.0)	26.3 (3.7)	27.3 (4.2)
Current smoker, %	25.1	23.0	11.3	18.5	23.0	20.0
Alcohol drinker, %	44.5	94.7	54.3	69.6	79.4	83.8
Prevalent cardiovascular disease, %	6.7	NA	0.0	6.5	8.1	5.1
Treated diabetes, %	8.5 <sup>*</sup>	1.9	25.8	7.9 <sup>**</sup>	10.6 <sup>**</sup>	10.9 <sup>**</sup>
Treated hypertension, %	27.2 <sup>*</sup>	4.2	34.9	16.9 <sup>**</sup>	32.5	28.6
Factor VII, n	7882	NA	3266	2801	1473	NA
Median factor VII activity, % (IQR)	116 (99-135)	NA	123 (107-143)	99 <sup>‡</sup> (89-110)	106 (93 - 120)	NA
Von Willebrand factor, n	8052	1461	NA	2804	3433	1846
Median von Willebrand factor antigen, % (IQR)	105 (81-134)	117.0 (93-145)	NA	121 (92-156)	125 (95 - 166)	114 (88 - 124)
Factor VIII, n	8047	NA	3226	NA	4006	NA
Median factor VIII activity, % (IQR)	121 (102-143)	NA	115 (95-140)	NA	112 (81 - 147)	NA

EA = European-ancestry participants; AA = African-ancestry participants; SD = standard deviation; NA = phenotype not available in cohort; IQR = interquartile range.

<sup>\*</sup> Includes those untreated and those self-reported at baseline.

<sup>\*\*</sup> Includes those untreated at baseline.

<sup>‡</sup> Antigen level in mg/dL.

Table 2

Description and association of loci associated with a single-nucleotide-polymorphism marker with p-values less than  $5.0 \times 10^{-8}$

Factor	Region	SNP			Variant	MAF	P-value	Parameter coefficient* (CI)	Closest Gene
		Number significant	Most significant	Position					
FVII, %	2p23	28	rs1260326	27584444	C→T	0.422	$6.2 \times 10^{-24}$	NE	0.0 kb from <i>GCKR</i> (intron)
	4q25	102	rs1126670	100271756	T→G	0.316	$3.6 \times 10^{-12}$	NE	0.0 kb from <i>ADH4</i> (exon)
	11q12	56	rs11230180	59718062	G→T	0.409	$2.0 \times 10^{-10}$	NE	11.5 kb from <i>MS4A6A</i>
	13q34	63	rs488703	112818877	C→T	0.116	$9.0 \times 10^{-259}$	NE	0.0 kb from <i>F7</i> (intron)
	20q11.2	56	rs867186	33228215	A→G	0.101	$5.7 \times 10^{-37}$	NE	0.0 kb from <i>PROCR</i> (exon)
vWF, %	6q24	79	rs9390459	147722052	G→A	0.442	$1.2 \times 10^{-22}$	-4.8 (-7.5, -2.1)	0.0 kb from <i>STXBP5</i> (exon)
	8p21	51	rs2726953	27857224	C→T	0.309	$1.3 \times 10^{-16}$	4.5 (1.5, 7.5)	0.0 kb from <i>SCARA5</i> (intron)
	9q34	178	rs687621	135126886	T→C	0.337	$<5.0 \times 10^{-324}$	24.1 (21.4, 26.8)	0.0 kb from <i>ABO</i> (intron)
	12p13	47	rs1063857	6023775	T→C	0.357	$1.7 \times 10^{-32}$	6.0 (3.2, 8.8)	0.0 kb from <i>VWF</i> (exon)
	12q23	2	rs4981022	102652341	T→C	0.315	$7.3 \times 10^{-10}$	-3.6 (-6.8, -0.4)	0.0 kb from <i>STAB2</i> (intron)
	12q24.3	20	rs7978987	129806374	G→A	0.346	$3.8 \times 10^{-11}$	3.4 (0.6, 6.2)	0.0 kb from <i>STX2</i> (intron)
	14q32	20	rs10133762	91362522	G→T	0.444	$2.3 \times 10^{-10}$	3.1 (0.4, 5.8)	0.0 kb from <i>TC2N</i> (intron)
FVIII, %	19p13.2	3	rs688875	7737166	A→G	0.262	$1.3 \times 10^{-9}$	-4.0 (-7.6, -0.4)	0.0 kb from <i>CLEC4M</i> (3' UTR)
	6q24	56	rs9390459	147722052	G→A	0.447	$6.7 \times 10^{-10}$	-2.6 (-4.9, -0.3)	0.0 kb from <i>STXBP5</i> (exon)
	8p21	26	rs9644133	27870402	C→T	0.192	$4.4 \times 10^{-15}$	-4.1 (-7.0, -1.2)	0.0 kb from <i>SCARA5</i> (intron)
	9q34	142	rs687289	133166660	C→T	0.343	$<5.0 \times 10^{-324}$	17.7 (15.4, 20.0)	0.0 kb from <i>ABO</i> (intron)
	12p13	3	rs1063856	6023795	A→G	0.354	$3.6 \times 10^{-9}$	2.6 (0.2, 5.0)	0.0 kb from <i>VWF</i> (exon)
	12q23	3	rs12229292	102656100	G→T	0.256	$7.2 \times 10^{-9}$	3.1 (0.2, 6.0)	0.0 kb from <i>STAB2</i> (intron)

MAF = weighted minor allele frequency; CI = confidence interval based on a 2-sided  $\alpha = 0.00000005$ ; NE = not estimated since antigen and antibody measure were combined; SNP = single nucleotide polymorphism.

\* Parameter coefficient represents change (% of activity or antigen) associated with 1-unit change in allele dosage.

Table 3

## a. Cohort specific results for the top loci associated with factor VII

CHR	SNP	Variant	ARIC		CHS		FHS		RS I	
			$\beta$	P-value	$\beta$	P-value	$\beta$	P-value	$\beta$	P-value
2	rs1260326	C→T	3.0	$8.0 \times 10^{-11}$	3.7	$7.2 \times 10^{-10}$	2.2	$4.6 \times 10^{-7}$	1.4	$1.6 \times 10^{-1}$
4	rs1126670	T→G	-3.1	$6.9 \times 10^{-11}$	-2.6	$6.1 \times 10^{-5}$	-0.1	$7.7 \times 10^{-1}$	-1.1	$3.0 \times 10^{-1}$
11	rs11230180	G→T	2.5	$5.5 \times 10^{-8}$	2.3	$3.7 \times 10^{-4}$	0.6	$2.5 \times 10^{-1}$	0.9	$3.3 \times 10^{-1}$
13	rs488703	C→T	-22.2	$5.7 \times 10^{-131}^{**}$	-16.9	$5.4 \times 10^{-106}$	-11.0	$2.2 \times 10^{-40}^{**}$	-15.1	$1.6 \times 10^{-21}$
20	rs867186	A→G	6.3	$1.3 \times 10^{-16}$	8.3	$5.9 \times 10^{-15}$	3.0	$1.0 \times 10^{-4}$	7.3	$6.3 \times 10^{-7}$

## b. Cohort specific results for the top loci associated with von Willebrand factor

CHR	SNP	Variant	ARIC		B58C		FHS		RS I		RS II	
			$\beta$	P-value	$\beta$	P-value	$\beta$	P-value	$\beta$	P-value	$\beta$	P-value
6	rs9390459	G→A	-4.1	$4.1 \times 10^{-10}$	-6.9	$4.0 \times 10^{-6}$	-4.9	$4.2 \times 10^{-5}$	-5.7	$1.5 \times 10^{-4}$	-5.2	$2.4 \times 10^{-3}$
8	rs2726953	C→T	5.0	$3.8 \times 10^{-12}^{*}$	1.1	$5.2 \times 10^{-1}^{*}$	5.1	$3.6 \times 10^{-4}^{*}$	3.4	$3.6 \times 10^{-2}^{*}$	5.8	$2.3 \times 10^{-3}^{*}$
9	rs687621	T→C	23.5	$6.0 \times 10^{-276}$	33.6	$4.0 \times 10^{-74}$	17.4	$7.0 \times 10^{-48}$	29.4	$1.8 \times 10^{-83}$	27.5	$7.9 \times 10^{-56}$
12	rs1063857	T→C	6.1	$4.1 \times 10^{-19}$	7.2	$3.5 \times 10^{-6}^{*}$	3.8	$2.2 \times 10^{-3}$	5.5	$4.0 \times 10^{-4}$	8.7	$6.6 \times 10^{-7}$
12	rs4981022	T→C	-3.4	$1.1 \times 10^{-5}^{*}$	-4.3	$2.7 \times 10^{-2}^{**}$	-1.1	$5.0 \times 10^{-1}^{**}$	-3.6	$2.3 \times 10^{-2}$	-6.9	$1.4 \times 10^{-4}$
12	rs7978987	G→A	3.9	$1.2 \times 10^{-8}$	2.1	$1.6 \times 10^{-1}$	2.9	$2.2 \times 10^{-2}$	3.2	$4.3 \times 10^{-2}$	2.7	$1.3 \times 10^{-1}$
14	rs10133762	G→T	3.2	$7.7 \times 10^{-7}$	5.2	$4.4 \times 10^{-4}$	2.5	$3.5 \times 10^{-2}$	0.9	$5.5 \times 10^{-1}$	3.5	$4.1 \times 10^{-2}$
19	rs868875	A→G	-3.8	$8.7 \times 10^{-6}^{**}$	-3.1	$3.9 \times 10^{-1}^{\dagger}$	-2.9	$2.8 \times 10^{-2}^{*}$	-5.1	$2.2 \times 10^{-2}^{**}$	-8.6	$9.5 \times 10^{-4}^{**}$

## c. Cohort specific results for the top loci associated with factor VIII

CHR	SNP	Variant	ARIC		CHS		RS I	
			$\beta$	P-value	$\beta$	P-value	$\beta$	P-value
6	rs9390459	G→A	-2.4	$9.1 \times 10^{-6}$	-3.1	$5.5 \times 10^{-4}$	-3.2	$4.9 \times 10^{-3}$
8	rs11780263	G→A	-4.2	$2.0 \times 10^{-10}$	-4.9	$7.1 \times 10^{-6}^{*}$	-2.6	$6.8 \times 10^{-2}$
9	rs687289	C→T	17.2	$4.1 \times 10^{-226}$	19.0	$3.0 \times 10^{-96}^{*}$	17.6	$1.4 \times 10^{-50}$
12	rs1063856	A→G	2.9	$1.6 \times 10^{-7}$	3.4	$2.4 \times 10^{-4}$	0.0	$9.8 \times 10^{-1}$



c. Cohort specific results for the top loci associated with factor VIII

CHR	SNP	Variant	ARIC		CHS		RS I	
			$\beta$	P-value	$\beta$	P-value	$\beta$	P-value
12	rs12229292	G→T	3.5	$2.1 \times 10^{-7}$ **	3.3	$2.5 \times 10^{-3}$ *	1.1	$4.2 \times 10^{-1}$

CHR = chromosome; P = p-value;  $\beta$  coefficient represents change (% of activity or antigen) associated with 1-unit change in allele dosage; SNP = single nucleotide polymorphism. All observed-expected variances (OEV) measuring imputation quality exceed 0.95 except those marked with

\* (OEV range from 0.80-0.94)

\*\* (OEV range from 0.50-0.79)

<sup>†</sup> (OEV <0.50).